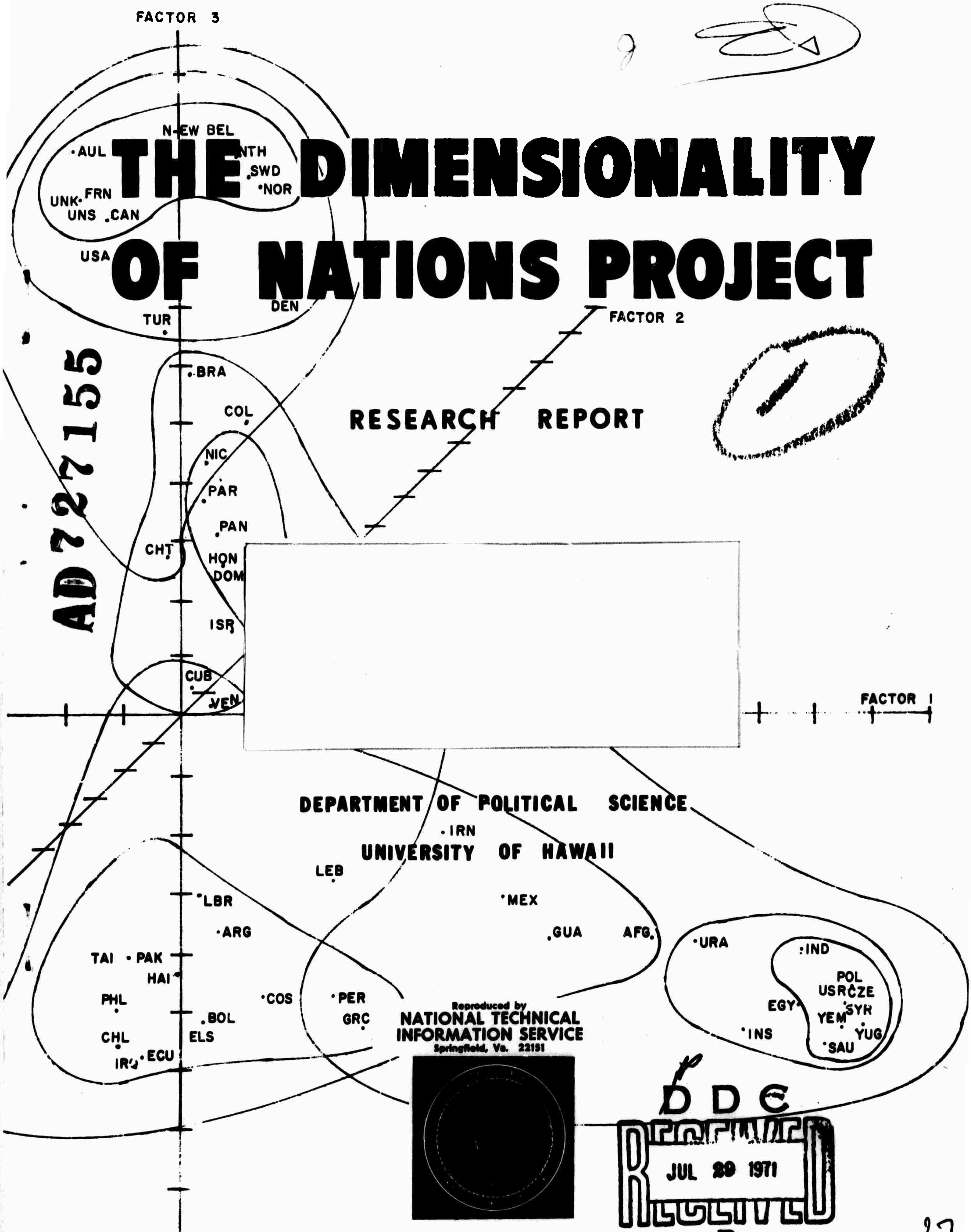
THE DIMENSIONALITY OF NATIONS PROJECT



RESEARCH REPORT

DEPARTMENT OF POLITICAL SCIENCE
UNIVERSITY OF HAWAII

27

DISCLAIMER NOTICE

THIS DOCUMENT IS THE BEST QUALITY AVAILABLE.

COPY FURNISHED CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

NOTICE TO USERS

Portions of this document have been judged by the NTIS to be of poor reproduction quality and not fully legible. However, in an effort to make as much information as possible available to the public, the NTIS sells this document with the understanding that if the user is not satisfied, the document may be returned for refund.

If you return this document, please include this notice together with the IBM order card (label) to:

National Technical Information Service
U.S. Department of Commerce
Attn: 952.12
Springfield, Virginia 22151

RESEARCH REPORT NO. 14

Computer Program Profile

Dennis R. Hall

JUN 1968

Prepared in Connection with the
National Science Foundation
Grant No. GS-1230
and the
Advanced Research Projects Agency
ARPA Order No. 1063
and Monitored by the
Office of Naval Research
Contract #N00014-67-A-0387-0003

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
The Dimensionality of Nations Project University of Hawaii 2500 Campus Road, Honolulu, Hawaii 96822	Unclassified 2b. GROUP

3. REPORT TITLE

Computer Program Profile

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report No. 14

5. AUTHOR(S) (First name, middle initial, last name)

Dennis R. Hall

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1968	24	0

8a. CONTRACT OR GRANT NO. N00014-67-A-0387-0003 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Research Report No. 14 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited and reproduction in whole or in part is permitted for any purpose of the United States government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
2500 Campus Road Honolulu, Hawaii 96822	Advanced Research Projects Agency Washington, D.C.

13. ABSTRACT

Computer techniques designed to construct numerical taxonomies from large, complex arrays of data have developed rapidly in the past few years. It has been the experience of many who have attempted to employ these techniques, however, that mathematically logical taxonomies were drawn. This difficulty has all too often left the researcher with a neat list of the sets of cases which are numerically similar, but with only the vaguest grasp of what the grouping similarities might be. This report introduces a simple computer technique to plot the underlying similarity of groups.

Unclassified

Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Computer Program						
National Profiles						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

Unclassified
Security Classification

TABLE OF CONTENTS

FIGURES

I INTRODUCTION

Computer techniques designed to construct numerical taxonomies from large, complex arrays of data have developed rapidly in the past few years. It has been the experience of many who have attempted to employ these techniques, however, that mathematically logical taxonomies are difficult to relate to the original data from which the taxonomies were drawn. This difficulty has all too often left the researcher with a neat list of the sets of cases which are numerically similar, but with only the vaguest grasp of what the grouping similarities might be. This report introduces a simple computer technique to plot the underlying similarity of groups.

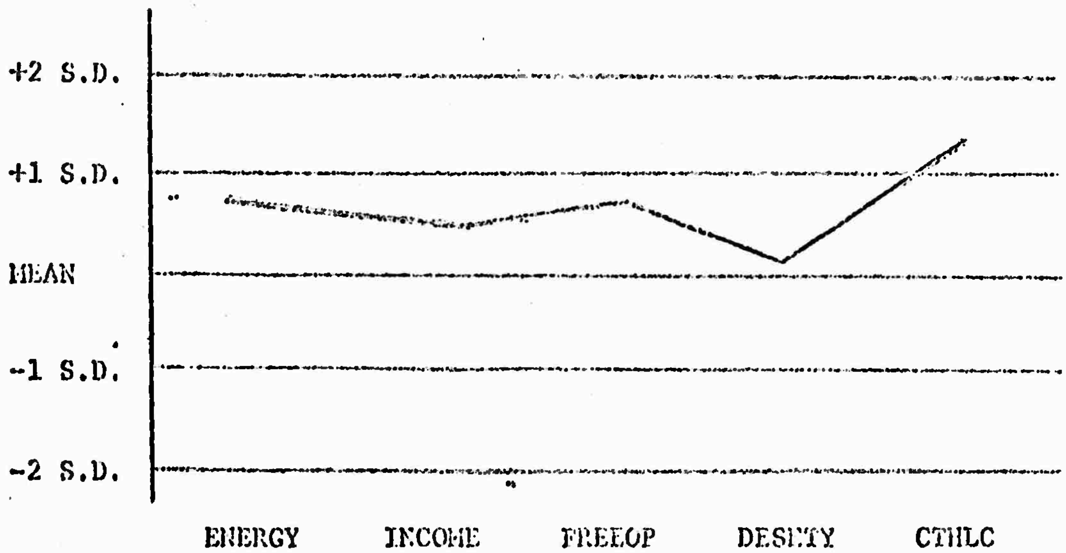
There are two basic types of numerical taxonomy; hierarchical, and cross-sectional. The former taxonomy -- the type generally employed in biology -- classifies life from a single instance of a living organism, through the species, phylum, kingdom, to all of organic existence. The hierarchical taxonomic scheme assumes <u>dynamic</u> similarity for grouping. A cross-sectional taxonomy, the type most generally employed by social scientists, assumes <u>static</u> similarity. A case is grouped as a member of one of k groups, each case is a member of one and only one group, and the group is assumed to be qualitatively distinct from all other groups.

II. DEVELOPMENT OF THE GROUPING PROBLEM

A Case is characterized by the values it has for a range of variables. These values may be specific, as the national income of France in U. S. dollars, or they may be general, as the freedom of group opposition in France measured by a two point scale. The numerical scores for these characteristics of the nation, France, form what we shall call an <u>attribute</u> (or characteristic) <u>profile</u> and can be shown graphically.

Figure 1

ATTRIBUTE PROFILE FOR FRANCE IN 1955*



*Plotted values are standard scores for France from five standardized characteristics on eighty-two nations. The plotted characteristics are:

ENERGY - energy consumption per capita

INCOME - national income

FREEOP - freedom of group opposition

DENSTY - population density

CATHLC - ratio of catholics to the population

To illustrate grouping procedure, let us assume that we are interested in constructing a taxonomy from eighty-two nations using as our index of similarity the five characteristics of Figure 1. If, on a transparent material, we were to draw eighty-two separate profile graphs, one for each nation, we could then build a taxonomy simply by superimposing sheets of the transparent material. Suppose we took the graph for France shown in Figure 1 and superimposed another graph from one of the eighty-one remaining transparent sheets, continuing the process until we found the line which coincided most closely with that for France. Let us assume that the second graph is the attribute profile of West Germany. Taking the graphs for both France and West Germany, we superimpose a third profile from among the remaining eighty graphs continuing until we again find the one most similar to the two already grouped. As we continued this exercise we would observe that the lines of our superimposed profile group would spread increasingly over the sheet. We would conclude that as the number of nations in the group increases, the similarity of profile for that group would decrease.

Let us assume that we were interested in building a hierarchical taxonomy from the transparent sheets. We would first lay the eighty-two profiles side by side and then find the two profiles out of all combinations of two which were most similar and superimpose them. This would leave eighty-one profiles. We would again look for the two profiles most similar and again superimpose them. As we continued this exercise we would find that the visual criteria for similarity of profile would have to be relaxed. We would continue to relax the similarity criteria until all eighty-two profiles were superimposed. If we kept account of the order in which the profiles were grouped, we could draw a taxonomic map of our procedure similar to that of Figure 2.

For social scientists, however, the hierarchy of profile groupings may not be salient and we might try to build a cross-sectional taxonomy from the eighty-two profiles assuming static similarity. We could start by specifying the k number of groups we were interested in or we could specify the level of group coherence or similarity we were willing to accept and then see how many groups resulted. Either way, we would try the various permutations of superimposition until we found the best cross-section for our purposes.

The permutations at each step of our hypothetical grouping procedures would be so numerous for the eighty-two profile graphs that they would preclude actually attempting to construct taxonomies in this way. We must look to computer techniques for assistance.

III. APPLICATION

The transparent graph illustration presents grouping procedure as the matching of characteristic profiles. The available computer techniques do not match profiles, but instead reduce the profile statistics to single indices of similarity or distance between the cases to be grouped, and then match the indices. The interpretation of computer taxonomies has been difficult because the indices of similarity upon which the methods depend reflect, but do not reproduce the original profile characteristics.

In the transparent graph illustration we had attempted to group cases by their scores on five uncorrelated variables. The standard scales by which the attribute profile values were measured can be viewed as Cartesian coordinates of a five dimensional space. Each of the coordinate axis of this space is at right angles to all others, since the variables are uncorrelated.It then becomes possible to represent the

nation profiles as single points in the Cartesian coordinate system of the five variables. Each nation has a unique location in this space and this location is a unique Euclidean distance from every other point in the space. The similarity (congruence) of profiles is thus measured by the similarity of spatial location -- the distance between points. The question of profile similarity may thus be reduced to measuring the Euclidean distances within this space.

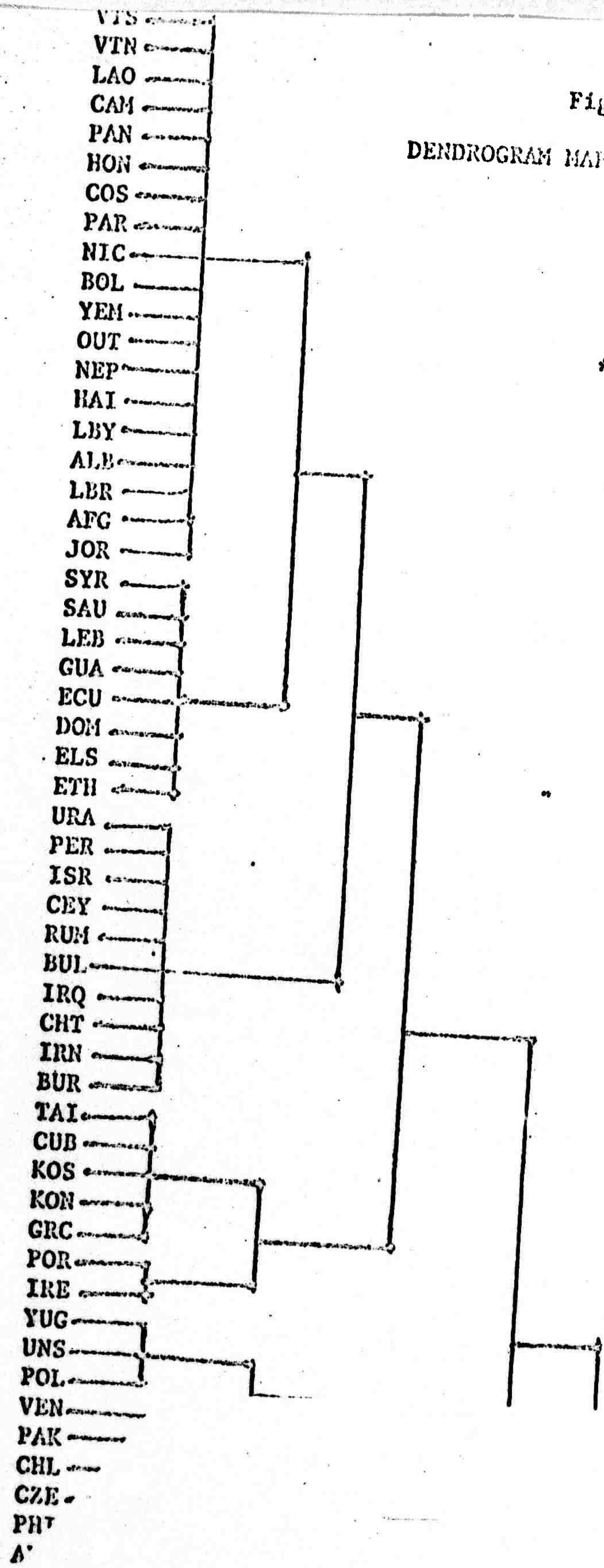
Since Euclidean distances between cases measure similarity, they can be used to develop taxonomies. Our concern at this point is not the development of a taxonomy, however, but with the underlying similarity the taxonomy represents.

A hierarchical taxonomic technique was employed to group the eighty-two nations on their Euclidean distances in the space of the five characteristics of Figure 1. A dendrogram displaying the resulting taxonomy is shown in Figure 2. The vertical lines represent the joining together of nations into a group. The distance between nations in the same group is increased as the procedure works from left to right in the dendrogram. At the extreme left no two nations are grouped and at the extreme right all nations form a single group.

As social scientists, we would probably not be interested in the entire dendrogram of Figure 2. Consequently, two interpretive problems arise. First, which groups should we extract from the dendrogram for presentation as our cross-section? Second, once we have selected our groups, how can we interpret the group similarities? To treat the latter problem we must return to the initial data - to the profiles themselves.

Figure 2

DENDROGRAM MAPPING OF GROUPS*



*Grouping of 82 nations based on their distances in five uncorrelated characteristics. Grouping method was S. C. Johnson, Hierarchical Clustering Scheme, Diameter Method.

IV. GENERAL DESCRIPTION

Assume that we are interested in characterizing a group of k size given us by the taxonomic method. We calculate a group score for each of the original characteristics by adding together the member scores and dividing by k, the number of members in the group. This mean score we will call a group profile score for a characteristic. Since we normally find variation about the group mean score (group dissimilarity) we can calculate the group standard deviation as a measure of member deviation from the group mean. This variance will vary from characteristic to characteristic and will serve as a measure of group cohesion on a characteristic. If we assume the distribution of member scores around the mean profile score for a group to be normal, we can add a confidence interval to the group mean score. A one standard deviation confidence interval about the group profile score would encompass approximately two-thirds of the member scores for the group across the characteristics. A group profile with its confidence interval can be shown graphically. (Figure 3).

The horizontal midpoint of the plot in Figure 3 is the population mean value for the characteristics. We would expect mean profile scores for any group we extract from a population to tend toward the horizontal midpoint of the plot since this portion of the graph is the most dense portion of the variable scatter. If the group mean profile score on any one of the characteristics is far removed from the midpoint, then that characteristic distinguishes the group: group members are similar on the characteristic. The tight cluster of group scores on national income (INCOME) in Figure 3 is taken as the distinguishing characteristic of the plotted group.

Figure 3

ATTRIBUTE PROFILE FOR GROUP FROM DENDROGRAM, FIG. 2*

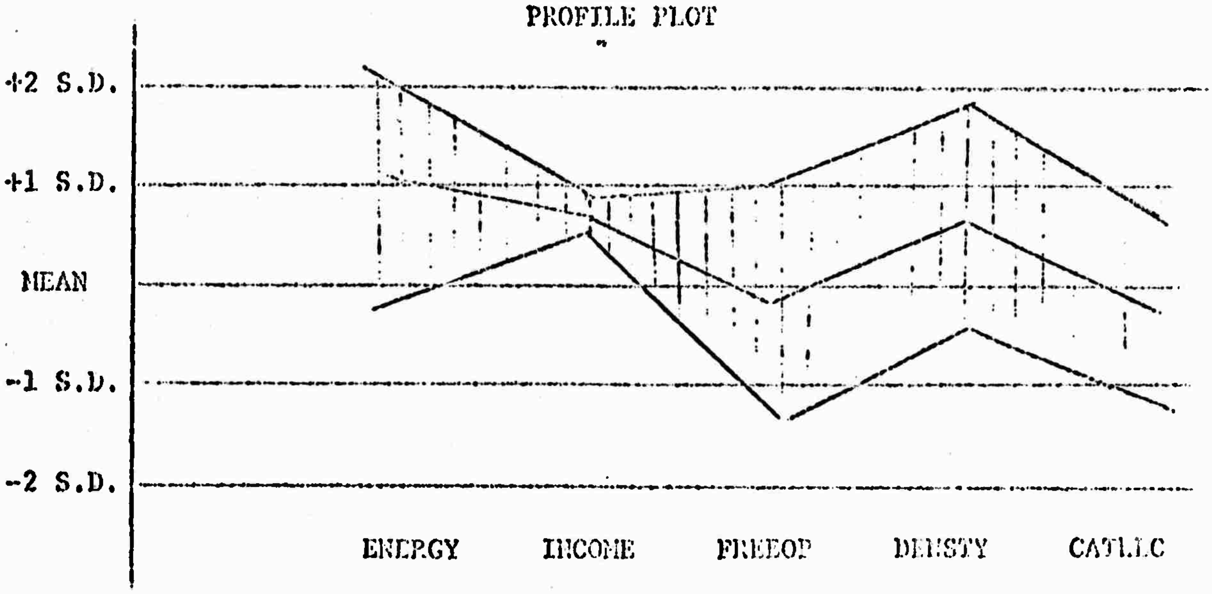
Members of the group are:

France China West Germany Russia Great Britain

Profile Statistics for Five Selected Characteristics

	ENERGY	INCOME	FREEOP	DENSTY	CATHLC
AVERAGE SCORE	1.03	0.78	-0.15	0.64	-0.20
STANDARD DEVIATION	1.104	0.136	1.144	1.097	0.864
GROUP RANGE	3.4	0.3	2.3	2.7	2.3

PROFILE PLOT



*Group was derived from S.C. Johnson's Hierarchical Clustering Scheme, Diameter Method. Ref. S.C. Johnson, "Hierarchical Clustering Schemes," *Psychometrika*, vol. 32, No. 3, September 1967.
ENERGY = energy consumption per capita; INCOME = national income;
FREEOP = freedom of group opposition; DENSTY = population density;
CATHLC = proportion of catholics to the population.

V. D.O.E. COMPUTER PROGRAM PROFILE WRITEUP

1. DESCRIPTION

A. This program calculates and plots group profiles.

B. Input (data) up to 200 cases, 50 variables.

C. Output

1. input matrix
2. means and standard deviations of input variables
3. standardized data matrix (option)
4. group profile statistics, up to 50 groups
5. group plots with one standard deviation confidence interval (option)

2. ORDER OF CONTROL CARDS

A. Plot Symbol Card

1. col. 1-62 b.X*-9-8-7-6-5-4-3-2-1b0b+1+2+3+4+5+6+7+8+9101112131415161718

B. Main Control Card

1. col. 1-6 PROBLM
2. col. 7-9 number of groups (maximum 50)
3. col. 10-12 number of cases (maximum 200)
4. col. 13-15 number of variables (maximum 50)
5. col. 18 1 standardize input data; blank do not standardize
6. col. 21 1 plot group profiles; blank do not plot

C. Group Size Card (maximum size of group 50 cases)

1. col. 1-3 number of cases in first group
2. col. 4-6 number of cases in second group
3. col. 7-9 number of cases in third group

. . . .

26. col. 76-78 number of cases in twenty-sixth group

(use a second card as needed to complete listing)

D. Observation Number Card (one set for each group)

1. col. 1-3 number of first case in the group
2. col. 4-6 number of second case in the group
3. col. 7-9 number of third case in the group

. . . .

26. col. 76-78 number of twenty-sixth case in the group (use a second card as needed to complete listing for the group)

E. Variable Format Card (use 2) both must be included even if second is blank.

F. Data Cards

G. Case Name Card

1. col. 1-6 name or number code of variable (as many cards as variables in the input data matrix. Blank cards must be inserted if variable names are not wanted)

I. Finish Card

1. col. 1-6 FINISH

3. MULTIPLE JOBS

A. For a multiple job, repeat card sets (A) through (H) for each job. Card (I) signals the completion of all jobs and is placed at the end of the final job.

```
      THIS PROGRAM STANDARDIZES INPUT, CALCULATES MEAN, RANGE, STANDARD PROF0010
      DEVIATION FOR GROUPS AND PLOTS THEIR PROFILE ACROSS THE VARIABLES PROF0020
      IT WILL ALSO CALCULATE THE PRODUCT-MOMENT CORRELATION  COEFFICIENTPROF0030
      BETWEEN GROUPS IT MAY BE REDIMENSIONED AS NEEDED                  PROF0040
                                                                        PROF0050
      DIMENSION NSIZE(50),A(200,50),NOBS(50,200),AVG(50,50),STDEV(50,50)PROF0060
     *,RANGE(50,50),VFMT(40),TMEAN(50),STAND(50),LINE(121),ZAVG(50,50),RPROF0070
     *(50,50),SYM(58)                                                   PROF0080
      DOUBLE PRECISION OBNAME(200),VNAME(50),NOB(50),FINISH,PROBLN      PROF0110
      REAL LINE                                                         PROF0100
      DATA FINISH/'FINISH'/                                             PROF0120
 10   FORMAT(A6,6I3)                                                    PROF0130
 11   FORMAT(26I3,2X)                                                   PROF0140
 12   FORMAT(20A4)                                                      PROF0150
 31   FORMAT(11X,121A1)                                                 PROF0160
 32   FORMAT(1X,I2,1X,A6,1X,121A1)                                      PROF0170
 60   FORMAT(1H1)                                                       PROF0180
 50   FORMAT(1X,I3,1X,A6,11F11.3)                                       PROF0190
 61   FORMAT(12X,'INPUT MATRIX A'/)                                     PROF0200
 62   FORMAT(12X,'STANDARDIZED DATA MATRIX B'/)                         PROF0210
 63   FORMAT(12X,'MEAN AND STANDARD DEVIATION OF INPUT VARIABLES'//)    PROF0220
 64   FORMAT(24X,'MEAN',12X,'STANDARD DEVIATION'//)                     PROF0230
 65   FORMAT(7X,I3,2X,A6,2X,F10.3,14X,F14.3/)                           PROF0240
 66   FORMAT(2X,'GROUP PROFILE VALUES AND STATISTICS')                  PROF0250
 67   FORMAT(/12X,'MEMBERS OF GROUP',I3)                                PROF0260
 68   FORMAT(/1X,'AVERAGE',4X,8(F10.4,3X))                              PROF0270
 69   FORMAT(/1X,'RANGE',6X,8(F10.4,3X))                                PROF0280
 70   FORMAT(/1X,'ST. DEV.',3X,8(F10.4,3X)/)                            PROF0290
 71   FORMAT(/17X,8(I3,10X))                                            PROF0300
 72   FORMAT(1H0,11(1X,I2,1X,A6,1X))                                    PROF0310
 73   FORMAT(/17X,11(I3,8X))                                            PROF0320
 74   FORMAT(1H1,'PROFILE PLOT OF GROUP',I2,1X,'WITH ONE STANDARD DEVIATPROF0330
     *ION CONFIDENCE INTERVAL'/)                                        PROF0340
 75   FORMAT(/10X,121A1)                                                PROF0350
 76   FORMAT(12X,'CHECK MEAN AND STANDARD DEVIATION OF STANDARDIZED MATRPROF0360
     *IX VARIABLES'//)                                                  PROF0370
 77   FORMAT(A6)                                                        PROF0380
 78   FORMAT(17X,11(A6,5X))                                             PROF0390
 80   FORMAT(2X/2X)                                                     PROF0400
 81   FORMAT(17X,8(A6,7X))                                              PROF0410
 90   FORMAT(63A1)                                                      PROF0420
 106  FORMAT(///1X,'THE NUMBER OF STANDARD DEVIATIONS FROM THE PLOT LEFTPROF0430
     * MARGIN TO THE ORIGIN IS',I2,'.  THE NUMBER OF SPACES PER STANDARDPROF0440
     *'/1X,'DEVIATION IS',I3,'.  THE ORIGIN IS LINE',I3,'.  THE RANGE OFPROF0450
     * DATA TO BE PLOTTED IS',F8.3,'.  THE MINIMUM VALUE IS',F7.3,'.')  PROF0460
 160  FORMAT(1H1,12X,'GROUP PRODUCT MOMENT CORRELATION COEFFICIENTS'////PROF0470
     *11X,'NO.',1X,16I7//)                                              PROF0480
 161  FORMAT(/6X,I6,4X,16F7.3)                                          PROF0490
```

```
C      READ PLOTTING SYMBOLS                                              PROF0530
       READ(5,90) BLANK,DOT,X,AST,(SYM(J),J=1,58)                         PROF0540
C      READ IN NUMBER OF GROUPS, NUMBER OF OBSERVATIONS, NUMBER OF        PROF0550
C      VARIABLES, AND OPTIONS TO STANDARDIZE INPUT, PLOT, AND CORRELATE   PROF0560
 409   READ(5,10)PROBLM,NGROUP,NCASE,NVAR,NSTAND,NPLOT,NCORR              PROF0570
       IF(PROBLM.EQ.FINISH) GO TO 9999                                    PROF0571
C      READ IN GROUP SIZES                                                PROF0580
       READ(5,11)(NSIZE(I),I=1,NGROUP)                                    PROF0590
C      READ IN NUMBER OF EACH OBSERVATION IN EACH GROUP                   PROF0600
       DO 1 I=1,NGROUP                                                    PROF0610
       L=NSIZE(I)                                                         PROF0620
       READ(5,11)(NOBS(I,J),J=1,L)                                        PROF0630
 1     CONTINUE                                                           PROF0640
C      READ VARIABLE FORMAT                                               PROF0650
       READ(5,12) VFMT                                                    PROF0660
C      READ INPUT DATA                                                    PROF0670
       DO2I=1,NCASE                                                       PROF0680
       READ(5,VFMT)(A(I,J),J=1,NVAR)                                      PROF0690
 2     CONTINUE                                                           PROF0700
       READ(5,77)(OBNAME(J),J=1,NCASE)                                    PROF0710
       READ(5,77)(VNAME(J),J=1,NVAR)                                      PROF0720
C      PRINT INPUT DATA                                                   PROF0730
       NMIN=0                                                             PROF0740
       I=(NCASE/51)+1                                                     PROF0750
       JMOD=MOD(NVAR,11)                                                  PROF0760
       IF(JMOD.EQ.0) J=(NVAR/11)                                          PROF0770
       IF(JMOD.NE.0) J=(NVAR/11)+1                                        PROF0780
       DO 402 II=1,I                                                      PROF0790
       NMIN=NMIN+50                                                       PROF0800
       DO 402 JJ =1,J                                                     PROF0810
       KBEG=(JJ-1)*11+1                                                   PROF0820
       NMIN=MINO(NCASE,NMIN)                                              PROF0830
       KEND=MINO((JJ*11),NVAR)                                            PROF0840
       WRITE(6,60)                                                        PROF0850
       WRITE(6,61)                                                        PROF0860
       IF(II.EQ.1) NBEG=1                                                 PROF0870
       IF(II.EQ.2) NBEG=51                                                PROF0880
       IF(II.EQ.3) NBEG=101                                               PROF0890
       IF(II.EQ.4) NBEG=151                                               PROF0900
       WRITE(6,73)(IND,IND=KBEG,KEND)                                     PROF0910
       WRITE(6,78)(VNAME(IND),IND=KBEG,KEND)                              PROF0920
       WRITE(6,80)                                                        PROF0930
       DO 403 KK=NBEG,NMIN                                                PROF0940
       WRITE(6,50) KK,OBNAME(KK),(A(KK,IND),IND=KBEG,KEND)                PROF0950
 403   CONTINUE                                                           PROF0960
 402   CONTINUE                                                           PROF0970
C      CALCULATE MEAN AND STANDARD DEVIATION OF INPUT DATA VARIABLES      PROF0980
       DO 7 J=1,NVAR                                                      PROF0990
```

```
      TOT=0.0
      TOTSQ=0.0
      DO 8 I=1,NCASE
      TOT=TOT+A(I,J)
      TOTSQ=TOTSQ+(A(I,J)**2)
 8    CONTINUE
      TMEAN(J)=TOT/FLOAT(NCASE)
      IF(((TOTSQ/FLOAT(NCASE))-(TMEAN(J)**2)).LE.0.0)  GO TO 87
      STAND(J)=SQRT((TOTSQ/FLOAT(NCASE))-(TMEAN(J)**2))
 87   CONTINUE
 7    CONTINUE
C     PRINT     MEAN AND STANDARD DEVIATION OF INPUT DATA VARIABLES
      WRITE(6,60)
      WRITE(6,63)
      WRITE(6,64)
      WRITE(6,65)(J,VNAME(J),TMEAN(J),STAND(J),J=1,NVAR)
C
C     TEST FOR STANDARDIZING INPUT DATA MATRIX OPTION
      IF(NSTAND.NE.1)  GO TO 99
C     STANDARDIZE  INPUT  VARIABLES
      DO 15 J=1,NVAR
      DO 9 I=1,NCASE
 9    A(I,J)=(A(I,J)-TMEAN(J))/STAND(J)
 15   CONTINUE
C     PRINT STANDARDIZED DATA MATRIX B
      NMIN = 0
      I=(NCASE/51)+1
      JMOD=MOD(NVAR,11)
      IF(JMOD.EQ.0) J=(NVAR/11)
      IF(JMOD.NE.0) J=(NVAR/11)+1
      DO 400 II=1,I
      NMIN=NMIN+50
      DO 400 JJ = 1,J
      KBEG=(JJ-1)*11+1
      NMIN=MIN0(NCASE,NMIN)
      KEND=MIN0((JJ*11),NVAR)
      WRITE(6,60)
      WRITE(6,62)
      IF(II.EQ.1) NBEG=1
      IF(II.EQ.2) NBEG=51
      IF(II.EQ.3) NBEG=101
      IF(II.EQ.4) NBEG=151
      WRITE(6,73)(IND,IND=KBEG,KEND)
      WRITE(6,78)(VNAME(IND),IND=KBEG,KEND)
      WRITE(6,80)
      DO 401 KK=NBEG,NMIN
      WRITE(6,59)KK,OBNAME(KK),(A(KK,IND),IND=KBEG,KEND)
 401  CONTINUE
```

```
 400  CONTINUE                                                          PROF1480
  99  CONTINUE                                                          PROF1490
C                                                                       PROF1500
C     CALCULATE GROUP MEAN, STANDARD DEVIATION, AND RANGE               PROF1510
      BMAX=-9999.0                                                      PROF1520
      BMIN=9999.0                                                       PROF1530
      DO 3 I = 1,NGROUP                                                 PROF1540
      DO 4 J = 1,NVAR                                                   PROF1550
      TOT=0.0                                                           PROF1560
      TOTSQ=0.0                                                         PROF1570
      AMAX=-9999.0                                                      PROF1580
      AMIN=9999.0                                                       PROF1590
      NSZ=NSIZE(I)                                                      PROF1600
      DO 5 K=1,NSZ                                                      PROF1610
      L = NOBS(I,K)                                                     PROF1620
      TOT=TOT+A(L,J)                                                    PROF1630
      TOTSQ=TOTSQ+(A(L,J)**2)                                           PROF1640
      IF(AMAX.LT.A(L,J)) AMAX = A(L,J)                                  PROF1650
      IF(AMIN.GT.A(L,J)) AMIN = A(L,J)                                  PROF1660
   5  CONTINUE                                                          PROF1670
      AVG(I,J) = TOT/FLOAT(NSIZE(I))                                    PROF1680
      STDEV(I,J)=0.0                                                    PROF1685
      IF((TOTSQ/FLOAT(NSIZE(I)))-(AVG(I,J)**2).LE.0.0) GO TO 85         PROF1690
      STDEV(I,J)=SQRT((TOTSQ/FLOAT(NSIZE(I)))-(AVG(I,J)**2))            PROF1700
  85  CONTINUE                                                          PROF1705
      IF(BMAX.LT.(AVG(I,J)+STDEV(I,J))) BMAX=AVG(I,J)+STDEV(I,J)        PROF1710
      IF(BMIN.GT.(AVG(I,J)-STDEV(I,J))) BMIN=AVG(I,J)-STDEV(I,J)        PROF1720
   4  RANGE(I,J) = AMAX-AMIN                                            PROF1740
   3  CONTINUE                                                          PROF1750
      RANG=BMAX+ABS(BMIN)                                               PROF1760
C     PRINT   GROUP MEAN, STANDARD DEVIATION, AND RANGE                 PROF1770
      WRITE(6,60)                                                       PROF1780
      WRITE(6,66)                                                       PROF1790
      DO 405 I=1,NGROUP                                                 PROF1800
      WRITE(6,67)     I                                                 PROF1810
      NSZ = NSIZE(I)                                                    PROF1820
      DO 6 J=1,NSZ                                                      PROF1830
      N=NOBS(I,J)                                                       PROF1840
   6  NOB(J)=OBNAME(N)                                                  PROF1850
      WRITE(6,72)(NOBS(I,J),NOB(J),J=1,NSZ)                             PROF1860
      JMOD=MOD(NVAR,8)                                                  PROF1870
      IF(JMOD.EQ.0) J = NVAR/8                                          PROF1880
      IF(JMOD.NE.0) J =(NVAR/8)+1                                       PROF1890
      DO 404 JJ = 1,J                                                   PROF1900
      KBEG=((JJ-1)*8)+1                                                 PROF1910
      KEND=MIN0((JJ*8),NVAR)                                            PROF1920
      WRITE(6,71)(N,N=KBEG,KEND)                                        PROF1930
      WRITE(6,81)(VNAME(N),N=KBEG,KEND)                                 PROF1940
```

```
      WRITE(6,68)(AVG(I,N),N=KBEG,KEND)                                 PROF19
      WRITE(6,70)(STDEV(I,N),N=KBEG,KEND)                               PROF19
      WRITE(6,69)(RANGE(I,N),N=KBEG,KEND)                               PROF19
  404 CONTINUE                                                          PROF19
      K=0                                                               PROF19
      DO 406 II=1,12                                                    PROF20
      K=K+4                                                             PROF20
      IF(I.NE.K) GO TO 406                                              PROF20
      WRITE(6,60)                                                       PROF20
  406 CONTINUE                                                          PROF20
  405 CONTINUE                                                          PROF20
C                                                                       PROF20
      IF(NPLOT.NE.1) GO TO 977                                          PROF2
C     PLOT GOUUP MEAN ACROSS VARIABLES WITH ONE S.DEV. CONFIDENCE       PROF24
C     INTERVAL ROUNDING THE PLOT LOCATIONS                              PROF24
      KSP=100.0/RANG                                                    PROF245
      LA=ABS(BMIN)+1.0                                                  PROF24
      KOR=(KSP*LA)+1                                                    PROF247
      WRITE(6,106) LA,KSP,KOR,RANG,BMIN                                 PROF24
      DO 999 I=1,NGROUP                                                 PROF249
      DO 105 J= 1,121                                                   PROF25
  105 LINE(J)=BLANK                                                     PROF251
      N=20-(2*LA)                                                       PROF252
      DO 51 II=1,121,KSP                                                PROF25
      M=N-1                                                             PROF25
      J=II+1                                                            PROF25
      LINE(II)=SYM(M)                                                   PROF25
      LINE(J)=SYM(N)                                                    PROF257
   51 N=N+2                                                             PROF25
      IF(LA.GT.9) LINE(KOR)=SYM(20)                                     PROF25
      WRITE(6,74) I                                                     PROF26
      NSZ=NSIZE(I)                                                      PROF261
      DO 52 J=1,NSZ                                                     PROF26
      N=NOBS(I,J)                                                       PROF26
   52 NOB(J)=OBNAME(N)                                                  PROF26
      WRITE(6,72)(NOBS(I,J),NOB(J),J=1,NSZ)                             PROF26
      WRITE(6,75) LINE                                                  PROF26
      DO 101 J=1,121                                                    PROF26
  101 LINE(J)= DOT                                                      PROF26
      WRITE(6,31) LINE                                                  PROF26
      DO 103 J=1,121                                                    PROF27
  103 LINE(J)= BLANK                                                    PROF27
      DO 190 NN=1,121,KSP                                               PROF27
  190 LINE(NN)=DOT                                                      PROF27
      DO 30 K=1,NVAR                                                    PROF27
      M=((AVG(I,K)*FLOAT(KSP))+0.5)+FLOAT(KOR)                          PROF27
      MR=(((AVG(I,K)+STDEV(I,K))*FLOAT(KSP))+0.5)+FLOAT(KOR)            PROF27
      MN=(((AVG(I,K)-STDEV(I,K))*FLOAT(KSP))+0.5)+FLOAT(KOR)            PROF27
```

```
      LINE(M)=X                                                    PROF2
      IF(MN.GT.120) GO TO 179                                      PROF2
      LINE(MN)=AST                                                 PROF2
  179 CONTINUE                                                     PROF2
      IF(MM.LT.0) GO TO 79                                         PROF2
      LINE(MM)=AST                                                 PROF2
   79 CONTINUE                                                     PROF2
      WRITE(6,32)K,VNAME(K), LINE                                  PROF2
      DO 104 J=1,121                                               PROF2
  104 LINE(J)=BLANK                                                PROF2
      DO 191 NN=1,121,KSP                                          PROF2
  191 LINE(NN)=DOT                                                 PROF2
      DO 988 JJ=1,4                                                PROF2
      WRITE(6,31) LINE                                             PROF2
  988 CONTINUE                                                     PROF2
   30 CONTINUE                                                     PROF2
  999 CONTINUE                                                     PROF2
  977 CONTINUE                                                     PROF2
      GO TO 409                                                    PROF2
 9999 CONTINUE                                                     PROF2
      STOP                                                         PROF2
      END                                                          PROF2
```

VI. SAMPLE PROBLEM

To characterize a cross-sectional taxonomy from the dendrogram of Figure 2, we selected a cross-section of eight groups and plotted their profiles with the profile program. The program features an automatic scaling device which scales the plots to the range of the statistics to be plotted in each job. The United States, the sole member of the eighth group, proved to be an extreme outlier on national income and energy consumption. To demonstrate the variable scale feature, groups 1 and 2 were plotted by themselves, without the necessity to scale the plots for the United States. The Plots for Groups 3 and 6 are products of the computer run when the United States was plotted. The computer output for these plots and statistics for the first three groups are shown on the following pages.

To allow for a large number of variables, plots are made vertically down the page, beginning with variable 1 at the top. The group mean score is denoted by an "X" and "*" denotes the one standard deviation confidence interval about the mean score. If a single asterisk appears, as in the group 1 plot for INCOME, member standard deviation is negligible and the confidence interval converges on a single point. If the asterisk is very close to the other asterisk, there may be no "X" between them. It is assumed that the mean score would fall between them in this case. The user may wish to shade the confidence intervals in the output as an aid in interpreting the profile variation across characteristics for the group.

GROUP PROFILE VALUES AND STATISTICS

MEMBERS OF GROUP 1

26 ETH	25 ILS	22 DOM	23 ECU	32 GLA	47 LEB	65 SAU	69 SYR	44 JOR	1 AFG
2 ALB	49 LBY	33 HAI	51 NEP	56 GUT	78 YEM	7 BUL	54 NIC	59 PAR	18 CGS
58 PAN	11 CAM	50 LAO	81 VTN	82 VTS	48 LBR	34 HON			

	1 ENERGY	2 INCOME	3 FREEDP	4 DENSTY	5 CATHLC
AVERAGE	-0.6257	-0.2433	-0.1250	-0.4162	0.0619
ST. DEV.	0.0777	0.0047	1.0218	0.5151	1.0753
RANGE	0.3160	0.0161	2.3359	1.8518	2.3565

MEMBERS OF GROUP 2

10 BUR	38 IRN	16 CHI	39 IRQ	9 BUL	64 RUM	15 CEY	41 ISR	60 PER	76 UKA

	1 ENERGY	2 INCOME	3 FREEDP	4 DENSTY	5 CATHLC
AVERAGE	-0.3730	-0.2259	-0.1510	0.0384	-0.4338
ST. DEV.	0.2590	0.0030	1.0182	0.9714	0.8283
RANGE	0.7443	0.0082	2.3359	3.2080	2.1423

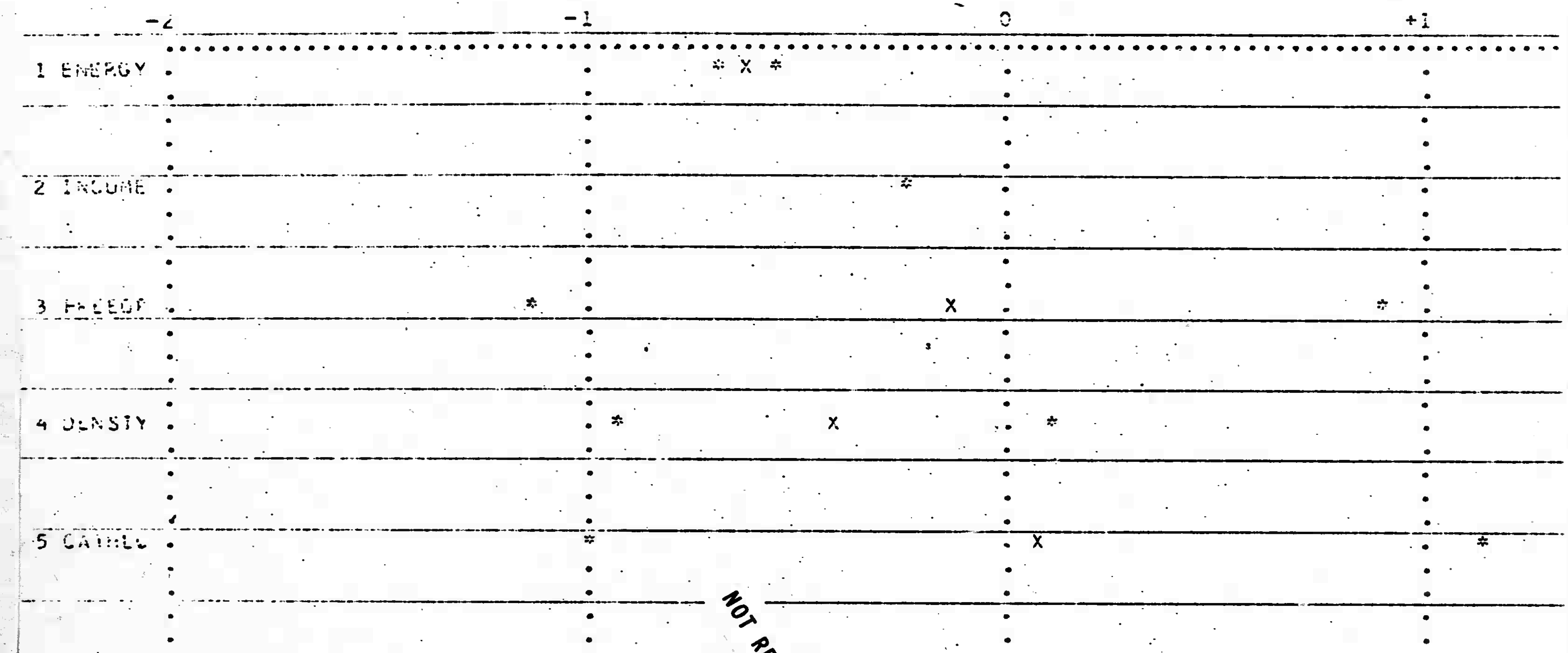
MEMBERS OF GROUP 3

40 INE	63 ROK	31 GRC	45 KON	46 KOS	19 CUB	70 TAI

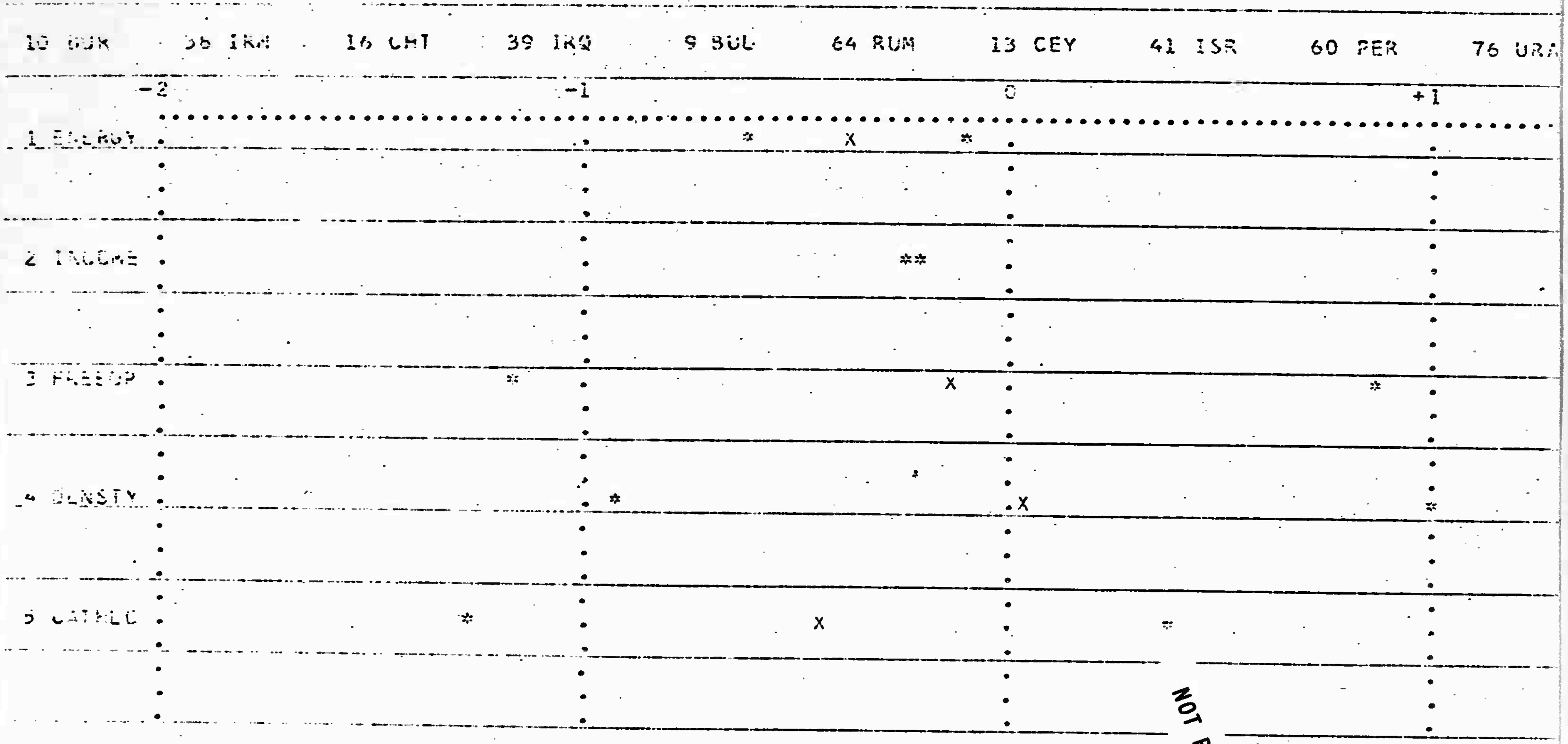
	1 ENERGY	2 INCOME	3 FREEDP	4 DENSTY	5 CATHLC
AVERAGE	[illegible]	-0.2041	-0.2177	[illegible]	0.0227
ST. DEV.	[illegible]	[illegible]	0.9729	[illegible]	1.1044
RANGE	[illegible]	[illegible]	2.3359	[illegible]	[illegible]

PROFILE PLOT OF GROUP 1 WITH ONE STANDARD DEVIATION CONFIDENCE INTERVAL .

26 ETH	25 ELS	22 DOM	23 ECU	32 GUA	47 LEB	65 SAU	69 SYR	44 JOR	1 AFG
2 ALB	49 LBY	33 HAI	51 NEP	56 OUT	78 YEM	7 BOL	54 NIC	59 PAR	18 COS
[illegible] PAN	11 CAM	[illegible] LAO	81 VTN	82 VTS	48 LBR	34 HON			

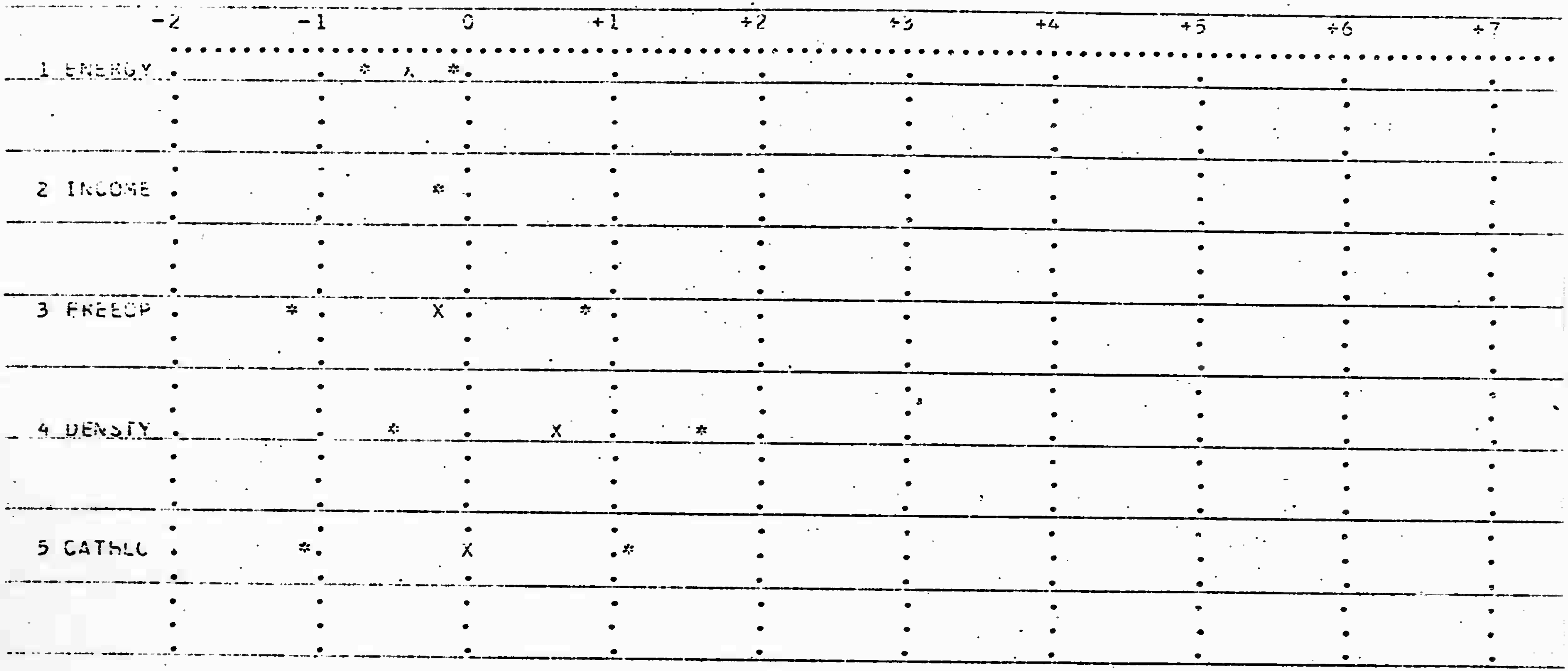


PROFILE PLOT OF GROUP 2 WITH ONE STANDARD DEVIATION CONFIDENCE INTERVAL



PROFILE PLOT OF GROUP 3 WITH ONE STANDARD DEVIATION CONFIDENCE INTERVAL

40 IRE 65 POR 31 GRC 45 KON 46 KOS 19 CUB 70 TAI



PROFILE PLOT OF GROUP 6 WITH ONE STANDARD DEVIATION CONFIDENCE INTERVAL